U.S. Geological Survey National Produced Waters Geochemical Database v2.2 (PROVISIONAL) Documentation

By Madalyn S. Blondes[1*], Kathleen D. Gans[2], James J. Thordsen[2], Mark E. Reidy[1], Mark A. Engle[1,3], Yousif K. Kharaka[2], Elizabeth L. Rowan[1], Burt Thomas[2,4]

[1]U.S. Geological Survey, Eastern Energy Resources Science Center, Reston, VA, USA

[2]U.S. Geological Survey, National Research Program, Menlo Park, CA, USA

[3]Dept. of Geological Sciences, University of Texas at El Paso, El Paso, TX, USA

[4]Now at: Dept. of Environmental and Earth Sciences, Willamette University, Salem, OR, USA

[*]Corresponding Author: Madalyn S. Blondes, email: mblondes@usgs.gov, telephone: (703) 648-6509

Disclaimers

Disclaimer for Provisional Database

The data you have secured from the U.S. Geological Survey (USGS) National Produced Waters Geochemical Database v2.2 are provisional and subject to revision. The data are released on the condition that neither the USGS nor the United States Government may be held liable for any damages resulting from their authorized or unauthorized use.

Distribution Liability

Although the data have been processed on computer systems at the USGS, U.S. Department of the Interior, no warranty, expressed or implied, is made by the U.S. Geological Survey regarding the utility of the data on any other system, nor shall the act of distribution constitute any such warranty. No responsibility is assumed by the USGS in the use of these data.

Additional Limitations

The information in the USGS National Produced Waters Geochemical Database v2.2 should be used with careful consideration of its limitations. The database is considered sufficiently accurate to provide an indication of tendencies in water composition from geographically and geologically defined areas. It is not appropriate for depiction of modern produced water compositions or examination of trends on small scales. The USGS makes no warranty regarding the accuracy or completeness of information presented in this database. Specific limitations of the database should be considered. Much of the information in the database cannot be independently verified. Methods of collection, sample preservation, analysis, assignment of geologic units and record keeping were not rigorous or standardized. Because of these uncertainties, users are advised to check data for inconsistencies,

outliers, and obviously flawed information. Methods of well construction, sample collection and chemical analysis have changed over time. Samples in the database may not be spatially representative and do not necessarily reflect the distribution and relative amount of water produced within a province and among geologic units. No sampling was planned to accurately depict the aggregate water composition of any area whether it be province, state, county or field. The geologic unit nomenclature developed for petroleum production may have changed over time. Water data collected 30 years ago from a province may not resemble water samples from current production. The composition of produced water within a province, field, or even well may change over time as a result of water flooding, recompletion in other intervals, and workovers. Water samples are commonly collected when a well has production problems or during the initial development of a well. Although criteria were applied to remove the obviously contaminated samples, the culling of unrepresentative data is considered incomplete. Most of the obvious redundant entries were removed from this database. Many of the remaining records represent multiple samples of the same well. Therefore aggregate statistics may be weighted by relatively few wells.

Version History

v2.2

Version 2.2 contains new input datasets and changes to the creation and format of the compiled database. New datasets, including KHARAKA, BARNABY, SKEEN, CAPO, ROWANMARCELLUS, HOBBS, KNOWLES, SHOUAKARSTASH, PERMIAN, ILLINOISSTEUBER, and MAFLA fill regional gaps and contain more isotopic information. The three original USGS datasets USGSMAIN, USGSOK, and USGSARK have been combined into a single one named USGSBREIT. Location errors were corrected and new well perforation depths were added from public sources for some data released in the previous version. Extraneous variables were removed and some variables were condensed into a single column. For example, all township and range information has been condensed into a single column TOWNRANGE. MGL and PPM have been condensed into UNITS. LAT, LATAPPROX, LONG, and LONGAPPROX have been condensed into LATITUDE and LONGITUDE, with a note in LATLONGAPX if the spatial data are approximate. All of the lithology columns were condensed into the single column LITHOLOGY. The database is now compiled using R (R Core Team, 2015) using a procedure described below. Variables with cleaned-up categorical data include WELLTYPE, STATE, USGSREGION, BASIN, ERA, and PERIOD. FORMATION names have been cleaned-up for only the Appalachian Basin. Duplicates are now found first within each input dataset and then tested between dataset using more stringent criteria. The new variable TDSUSGS is a reported or calculated value in mg/L for ease of plotting. The new variable SG is a reported or empirically determined specific gravity that allows the user to convert from ppm to mg/L.

v2.1

Version 2.1 corrects errors found in version 2.0 of the database. Incorrect LATITUDE, LONGITUDE, or STATE variables were updated based on API or other well information. Chemical and well data in incorrect columns were placed in the correct columns. Unit problems were fixed for chemistry and specific gravity data. Alkalinity data were put into the correct columns based on the method of measurement. Certain variables not given in the original input datasets, including WELLTYPE and reservoir age information were determined based on well and formation data. Various other errors noted by users were corrected by referring back to the original source of the data. No new datasets were added except IDDB = "WILLISTON," which is a compilation of the EASTPOPLAR and BAKKEN entries from version 2.0 of the database along with unpublished data (Thamke, 2014, written communication).

Introduction

During hydrocarbon production, water is typically co-produced from the geologic formations producing oil and gas. Understanding the composition of these produced waters is important to help investigate the regional hydrogeology, the source of the water, the efficacy of water treatment and disposal plans, potential economic benefits of mineral commodities in the fluids, and the safety of potential sources of drinking or agricultural water. In addition to waters co-produced with hydrocarbons, geothermal development or exploration brings deep formation waters to the surface for possible sampling. This U.S. Geological Survey (USGS) Produced Waters Geochemical Database, which contains geochemical and other information for 166,994 produced water and other deep formation water samples of the United States, is a provisional, updated version of the 2002 USGS Produced Waters Database (Breit and others, 2002). In addition to the major element data presented in the original, the new database contains trace elements, isotopes, and time-series data, as well as nearly 100,000 additional samples that provide greater spatial coverage from both conventional and unconventional reservoir types, including geothermal. The database is a compilation of 33 individual databases, publications, or reports. The database was created in a manner to facilitate addition of new data and correct any compilation errors, and is expected to be updated over time with new data as provided and needed. Table 1 shows the abbreviated ID of each input database (IDDB), the number of samples from each, and its reference. Table 2 defines the 192 variables contained in the database and their descriptions. The database variables are organized first with identification and location information, followed by well descriptions, dates, rock properties, physical properties of the water, and then chemistry. The chemistry is organized alphabetically by elemental symbol. Each element is followed by any associated compounds (e.g. H_2S is found after S). After Zr, molecules containing carbon follow,

including measures of alkalinity, dissolved organic carbon (DOC), and hydrocarbons. Isotopic data are found at the end of the dataset, just before the culling parameters.

Database Compilation Procedure

Input Datasets

Modification of the original data or variable names is necessary to create a database with consistent headers, compositional units, and numeric data that can be plotted or analyzed as a whole. One of the main goals of this updated database is to create a compiled dataset where every change to the original datasets is reversible and recorded. Thus if errors are found, there is a coded record that can be adjusted as needed, and the compiled dataset can be easily recreated from the original data files. To meet this goal, the USGS National Produced Waters Geochemical Database v2.2 is compiled using R, a language and environment for statistical computing and graphics (R Core Team, 2015)[1]. An R script is written for each input database that imports the original data, renames the variables to match the template (Table 2), and then appends the existing columns to a template header. Non-numeric characters within numeric variables (for example, ions and pH) are fixed, deleted, or replaced with the following codes:

- <MDL = Less than the Method Detection Limit (e.g. N.D., non-detect, absent)
- <PQL = Less than the Practical Quantification Limit (e.g. Trace, minor, present)
- UNK = Transcription error or otherwise nonsensical entry (e.g. 10K41)
- NA = Not analyzed, unknown (e.g. --, n.a., NA)

[1] Disclaimer: Use of brand or trade names are for descriptive purpose and do not imply endorsement by the U.S. Government

Censored data for which the MDL is known are left as is (e.g. <10 remains <10 rather than <MDL). Dates are formatted into a consistent date form and extra variables are removed. Units for all variables other than the major and minor ions are defined in Table 2. The major and minor ions are generally reported in units of milligrams per liter (mg/L) or on a mass basis as parts per million (ppm) or its equivalent, milligrams per kilogram (mg/kg). The units used are defined in the UNITS column. The user of this database must be careful to examine the units when using the data and can convert between the two using measurements or estimates of brine density. A calculated TDS variable (TDSUSGS) is described below and has already been converted to mg/L for easy plotting. This variable is different than TDSCALC, which is the reported value calculated by the author of an input dataset or report.

Samples with duplicate American Petroleum Institute well identification (API) codes and chemical concentrations with large numbers of significant figures in common, were culled. It is unlikely that two samples, even from the same well, will have the exact same concentrations for three or more analytes, and therefore such similarities represent true duplicates. API code, calcium (Ca), chloride (Cl), and bicarbonate (HCO_3) concentrations are used for a starting duplicate search. If these analytes were detected in samples that did not appear to be true duplicates, other analytes like Na and Mg were added to the search. Care was taken to avoid false duplicates (e.g. where all three of the initial ions had null data).

Once non-numeric characters were replaced and duplicates were removed, categorical and text variables were cleaned up. Variables such as SOURCE, REFERENCE, WELLTYPE and STATE were defined. Lastly, the new dataset was saved as an intermediate file to be appended to the main database.

Compiled Dataset

Each individual input database is appended to the template using a global R script. The database is further standardized here with:

1. Internally consistent 14-digit API (API)
2. One of eight regions (USGSREGION) defined by the U.S. Geological Survey National Oil and Gas Resource Assessment Team (1995)
3. Basins simplified to 85 different reported names (BASIN)
4. State names (STATE)
5. One of eight well type (WELLTYPE) designations (Conventional Hydrocarbon, Shale Gas, Tight Oil, Tight Gas, Coal Bed Methane, Geothermal, and Groundwater)
6. Geologic Era (ERA)
7. Geologic Period (PERIOD)
8. Formations (FORMATION) have been standardized only for data in the Appalachian Basin.

Duplicates are again removed between input datasets using similar criteria as described above. The duplicate observation retained was generally the one in the database that contained more information.

Calculated Variables

The new column SG provides an option for conversion from ppm to mg/L. SG is either the reported specific gravity, or an estimate based on a relationship between the reported specific gravity and the total dissolved solids (TDS) measured in ppm for the entire database. For example, if calcium (Ca) concentrations are in ppm, it simply can be multiplied by SG to obtain Ca concentration in mg/L. This has already been done for the new column TDSUSGS. The intent of the TDSUSGS column is to

provide the user with an easy way to plot the most common variable with consistent units. TDSUSGS shows the TDS concentration, in mg/L, in the following order of primacy:

1. the reported measured TDS
2. the reported calculated TDS
3. the sum of all major cations and anions assuming both Na and Cl data exist. TDS concentrations in ppm were converted to mg/L by multiplying by SG.

Culling Data Based on Chemistry

Quality control of the dataset can be performed by culling based on geochemical criteria. In this version 2.2 of the provisional database, the data that fall outside of the bounds of the following criteria are flagged, rather than culled. There are six temporary columns in the database that indicate the failure to meet specific culling criteria, based on those published in Hitchon and Brulotte (1994). An "X" is placed in the columns shown in Table 2 where the sample falls outside of the pH range of 4.5 – 10.5; where Mg > Ca, K > Cl, or K > 5 x Na; and where the charge balance error is greater than 15%. The data were kept to allow the user to make judgment calls on their quality.

Files Available for Download

The database is available in three different file formats: a comma separated values (.csv) text file, a Microsoft Excel spreadsheet (.xlsx) and an .Rdata file for R users. There are two versions for each of these, one with a "c" suffix and one with an "n" suffix. The "c" files retain all text codes within the numeric variables that describe the data (e.g. "<0.01", "<MDL", "<PQL", or "UNK"). The "n" files remove all non-numeric codes from the numeric variables for easy plotting. The following six files are available for download on the USGS Produced Waters website:

- USGS_Produced_Waters_v2.2c.csv
- USGS_Produced_Waters_v2.2c.xlsx
- USGS_Produced_Waters_v2.2c.Rdata
- USGS_Produced_Waters_v2.2n.csv
- USGS_Produced_Waters_v2.2n.xlsx
- USGS_Produced_Waters_v2.2n.Rdata

References Cited

Barnaby, R.J., Oetting, G.C., and Gao, G., 2004, Strontium isotopic signatures of oil-field waters: Applications for reservoir characterization: AAPG Bulletin, v. 88, no. 12, p. 1677–1704.

Barton, G.J., Burruss, R.C., and Ryder, R.T., 1998, Water quality in the vicinity of Mosquito Creek Lake, Trumbull County, Ohio, in relation to the chemistry of locally occurring oil, natural gas, and brine: U.S. Geological Survey Water-Resources Investigation Report 98-4180, 46 p.

Bassett, R.L., and Bentley, M.E., 1983, Deep brine aquifers in the Palo Duro Basin: regional flow and geochemical constraints: Bureau of Economic Geology, University of Texas, Report of Investigations, no. 130.

Bownocker, J.A., 1906, Salt deposits and the salt industry in Ohio: Ohio Division of Geological Survey Bulletin 8, 42 p.

Breen, K.J., Angelo, C.G., Masters, R.W., and Sedam, A.C., 1985, Chemical and isotopic characteristics of brines from three oil- and gas-producing sandstones in eastern Ohio, with applications to the geochemical tracing of brine sources: U.S. Geological Survey Water-Resources Investigations Report WRI 84-4314, 58 p., Accessed *http://pubs.er.usgs.gov/djvu/WRI/wrir_84_4314.djvu*.

Breit G.N., 2002, USGS Produced Waters Database, Accessed *http://energy.cr.usgs.gov/prov/prodwat.*

Capo, R.C., Stewart, B.W., Rowan, E.L., Kolesar Kohl, C.A., Wall, A.J., Chapman, E.C., Hammack, R.W., and Schroeder, K.T., 2014, The strontium isotopic evolution of Marcellus Formation produced waters, southwestern Pennsylvania: International Journal of Coal Geology, v. 126, p. 57–63.

Carpenter, A.B., Trout, M.L., and Picket, E.E., 1974, Preliminary report on the origin and chemical evolution of lead and zinc-rich oil field brines in central Mississippi: Economic Geology, 1974, v. 69, p. 1191-1206.

Cimarex Energy Company, 2013, written communication.

Clifford, M.J., 1975, Subsurface liquid-waste injection in Ohio: Ohio Division of Geological Survey Information Circular 43, 27 p.

Conrey, G.W., 1921, The geology of Wayne County: Ohio Division of Geological Survey Bulletin 24, 155 p.

Dahm, K.G., 2013, written communication: U.S. Bureau of Reclamation.

Dahm, K.G. and others, 2011, Composite geochemical database for coalbed methane produced water quality in the Rocky Mountain region: Environmental Science and Technology, v. 45, p. 7655-7663.

Department of Energy, National Energy Technology Laboratory, NATCARB Brine Database, Accessed 2013 at www.netl.doe.gov.

Department of Energy, National Energy Technology Laboratory, Rocky Mountain Produced Waters Database, 2005, Accessed http://www.alrc.doe.gov/technologies/oil-gas/Software/database.html.

Dresel, P. E., and Rose, A. W., 2010, Chemistry and origin of oil and gas well brines in western Pennsylvania: Pennsylvania Geological Survey,. 4th series, Open-File Report OFOG 10–01.0, 48 p., web version, accessed on December 28, 2012, Accessed *http://www.dcnr.state.pa.us/topogeo/pub/openfile/pdfs/ofog10_01.pdf.*

Hanshaw, B.B., and Hill, G.A., 1969, Geochemistry and hydrodynamics of the Paradox Basin region, Utah, Colorado, and New Mexico: Chemical Geology, v. 4, p. 263-164.

Hayes, T., 2009, Sampling and analysis of water streams associated with development of Marcellus Shale gas, Final Report, December 31, 2009: Marcellus Shale Coalition, p. 107-356, Accessed http://catskillcitizens.org/learnmore/20100916IOGAResponsetoDECChesapeake_IOGAResponsetoDEC.pdf.

Hitchon, B., and Brulotte, M., 1994, Culling criteria for “standard” formation water analyses: Applied Geochemistry, v. 9, p. 637-645.

Hobbs, M.Y., Frape, S.K., Shouakar-Stash, O., and Kennell, L.R., 2011, Regional Hydrogeochemistry, Southern Ontario: Nuclear Waste Management Organization OPG's Deep Geologic Repository for low and intermediate level waste NWMO DGR-TR-2011-12, 157 p.

Hoskins, H.A., 1947, Analyses of West Virginia brines: West Virginia Geological and Economic Survey, Report of Investigations, no. 1, p. 1-21.

Keller, S.J., 1983, Analyses of subsurface brines of Indiana: Department of Natural Resources Geological Survey Occasional Paper 41.

Kharaka, Y.K., 1971, Simultaneous flow of water and solutes through geological membranes: experimental and field investigations: University of California, Berkeley.

Kharaka, Y.K., and Berry, F.A., 1974, The influence of geological membranes on the geochemistry of subsurface waters from Miocene sediments at Kettleman North Dome in California: Water Resources Research, v. 10, no. 2, p. 313–327.

Kharaka, Y.K., Berry, F.A., and Friedman, I., 1973, Isotopic composition of oil-field brines from Kettleman North Dome, California, and their geologic implications: Geochimica et Cosmochimica Acta, v. 37, no. 8, p. 1899–1908.

Kharaka, Y.K., Callender, E., and Carothers, W.W., 1977, Geochemistry of geopressured geothermal waters from the Texas Gulf Coast: University of Southwestern Louisiana, Novemeber 16 - 18 Proceedings Third Geopressured Geothermal Energy Conference, v. 1, GI–121 – GI–166 p.

Kharaka, Y.K., Leong, L.Y.C., Doran, G., and Breit, G.N., 1998, Can produced water be Reclaimed? Experience with placerita oil field, California, *in* 5th International Petroleum Environmental Conference – Proceedings of the 5th International Petroleum Environmental Conference: Albuquerque, N.M, Oct. 20-23, p. 1145–1169.

Kharaka, Y.K., Maest, A.S., Carothers, W.W., Law, L.M., Lamothe, P.J., and Fries, T.L., 1987, Geochemistry of metal-rich brines from central Mississippi Salt Dome basin, USA: Applied Geochemistry, v. 2, no. 5, p. 543–561.

Kharaka, Y.K., and Otton, J.K., 2003, Environmental impacts of petroleum production: initial results from the Osage-skiatook petroleum environmental research sites, Osage county, Oklahoma: US Department of the Interior, US Geological Survey Oklahoma, OK U.S. Geological Survey Water-Resources Investigations 03-4260, 159 p.

Kharaka, Y.K., and Thordsen, J.J., 1992, Stable isotope geochemistry and origin of waters in sedimentary basins, *in* Isotopic signatures and sedimentary records: Springer, p. 411–466.

Kharaka, Y.K., and Carothers, W.W., 1988, Geochemistry of oil-field water from the North Slope, *in* Gryc, George ed., Geology and exploration of the National Petroleum Reserve in Alaska, 1974 to 1982:, p. 551–561.

Kharaka, Y.K., and Hanor, J.S., 2014, Deep Fluids in Sedimentary Basins, *in* Holland H.D. and Turekian, K.K. eds., Treatise on Geochemistry: Elsevier: Oxford, p. 472–515.

Lamborn, R.E., 1952, Additional analyses of brines from Ohio: Ohio Division of Geological Survey Report of Investigations 11, 56 p.

Lawhead, A.H., no date, Natural brines in the Illuminating Company service area: unpublished, Cleveland Electric Illuminating Company report, not paged.

Leahy, J.N., 1996, Abandoned Alpha Portland Cement Company mine facility site, Ironton, Ohio: a groundwater flow and brine migration analysis: M.S. Thesis, University of Cincinnati, 246 p.

Lowry, R. M., Faure, G., Mullet, D. I., and Jones, L. M., 1988, Interpretation of chemical and isotopic compositions of brines based on mixing and dilution, "Clinton" sandstones, eastern Ohio, U.S.A.: Applied Geochemistry, v. 3, p. 177-184.

McDonald, J., Riley, R.A., Wolfe, M.E., Wells, J.G., 2013, written communication, Brine geochemistry database of Ohio: Ohio Division of Geological Survey.

McGrain, P., 1953, Miscellaneous analyses of Kentucky brines: Kentucky Geological Survey- Report of Investigations No. 7, p. 1-16.

McGrain, P., and Thomas, G. R., 1951, Preliminary report on the natural brines of eastern Kentucky: Kentucky Geological Survey, v. Report of Investigations No. 3, p. 1-22.

Meents, W.F., Bell, A.H., Rees, O.W., and Tilbury, W.G., 1952, Illinois oil-field brines Their geologic occurrence and chemical composition: Illinois State Geological Survey, Illinois Petroleum no. 66, 38 p.

Meredith, E., Kuzara, S., Wheaton, J.R., Bierbach, S., Chandler, K., Donato, T., Gunderson, .J., and Schwartz, C., 2011, 2010 Annual coalbed methane regional groundwater monitoring report: Powder River Basin, Montana, Montana Bureau of Mines and Geology Open File Report 600, 130 p., 6 sheets, Accessed http://www.mbmg.mtech.edu/pdf-open-files/mbmg600-2010CBM-Report.pdf.

Mills, R. V. A., and Wells, R. C., 1919, The evaporation and concentration of waters associated with petroleum and natural gas: U.S. Geological Survey Bulletin 693, 104 p.

Moldovanyi, E.P. and Walter, L.M., 1992, Regional trends in water chemistry, Smackover Formation, southwest Arkansas: Geochemical and Physical Controls, AAPG Bulletin, v. 76, no. 6, p 864-894.

New Mexico Water and Information Data System (NM WAIDS), Accessed 2013 at http://octane.nmt.edu/waterquality.

North Dakota Oil and Gas Division, Accessed 2013 at https://www.dmr.nd.gov/oilgas/.

New York State Department of Environmental Conservation, 2009, Draft supplemental generic environmental impact statement on the oil, gas, and solution mining regulatory program, well permit issuance for horizontal drilling and high-volume hydraulic fracturing to develop the Marcellus Shale

and other low-permeability gas reservoirs: New York State Department of Environmental Conservation (DEC), Appendix 13, Accessed http://www.dec.ny.gov/energy/58440.html.

Orton, Edward, 1888, The Trenton Limestone as a source of oil and gas in Ohio: Ohio Division of Geological Survey, vol. 6, p. 101-310.

Osborn, S.G., and McIntosh, J.C., 2010, Chemical and isotopic tracers of the contribution of microbial gas in Devonian organic-rich shales and reservoir sandstones, northern Appalachian Basin: Applied Geochemistry, v. 25, p. 456-471.

Osborn, S.G., McIntosh, J.C., Hanor, J.S., and Biddulph, D., 2012, Iodine-129, 87Sr/86Sr, and trace elemental geochemistry of northern Appalachian Basin brines: evidence for basinal-scale fluid migration and clay mineral diagenesis: American Journal of Science, v. 312, p. 263-287.

Pashin, J., 2013, written communication.

Pennsylvania Department of Environmental Protection, Bureau of Oil and Gas Management, 1991, NORM Survey Summary, Accessed http://files.dep.state.pa.us/OilGas/BOGM/BOGMPortalFiles/RadiationProtection/NORM.pdf.

Peterman, Z.E., Thamke, J.N., Futa, K., Oliver, T.A., 2010, Strontium isotope detection of brine contamination in the East Poplar Oil Field, Montana: U.S. Geological Survey Open-File Report 2010-1326, 20 p., accessed http://pubs.usgs.gov/of/2010/1326/.

Poth, C.W., 1962, The occurrence of brine in western Pennsylvania: Pennsylvania Geological Survey, v. Fourth Series, Bulletin M 47, 53 p.

Powell, W.J., and Carroon, L.E., 1963, Water problems associated with oil production in Alabama: Geological Survey of Alabama, Division of Water Resources, Circular 22, p. 63.

Preston, T.M., Smith, B.D., Thamke, J.N., Chesley-Preston, T., 2012, Water-quality and geophysical data for three study sites within the Williston Basin and Prairie Pothole Region: U.S. Geological Survey Open-File Report 2012-1149, 17 p., accessed http://pubs.usgs.gov/of/2012/1149/.

Price, P.H., Hare, C.E., McCue, J.B., and Hoskins, H.A., 1937, Salt Brines of West Virginia: West Virginia Geological Survey, v. VIII, 203 p.

R Core Team, 2015, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Ransome, F.L., 1913, Contributions to economic geology, 1911, Part I, Metals and nonmetals except fuels--Copper; Lead and Zinc: U.S. Geological Survey Bulletin 530-B, 400 p.

Reeves, F., 1917, The absence of water in certain sandstones of the Appalachian oil fields: Economic Geology, v. 12, p. 354-378. [PA and WV]

Rice, C.A., 2003, Production waters associated with the Ferron coalbed methane fields, central Utah: Chemical and isotopic composition and volumes, International Journal of Coal Geology, 56: 141-169.

Rice, C.A., and others, 2000, Water co-produced with coalbed methane in the Powder River Basin, Wyoming: preliminary compositional data, United States Geological Survey, Open File Report 00-372.

Root, W.J., 1888, The manufacture of salt and bromine: Ohio Division of Geological Survey, vol. 6, p. 653-670.

Rowan, E.L., Engle, M.A., Kirby, C., and Kraemer, T.F., 2011, Radium content of oil- and gas-field produced waters in the northern Appalachian Basin: Summary and discussion of data: U.S. Geological Survey, Scientific Investigations Report 2011-5135, 31 p.

Rowan, E.L., Engle, M.A., Kraemer, T.F., Schroeder, K.T., Hammack, R.W., and Doughten, M.W., 2015, Geochemical and isotopic evolution of water produced from Middle Devonian Marcellus shale gas wells, Appalachian basin, Pennsylvania: AAPG Bulletin, v. 99, no. 02, p. 181–206.

Sanders, L.L., 1986, Geochemistry and paleotemperature of formation waters from the Lower Silurian "Clinton" formation, Eastern Ohio: Ph.D. Dissertation, Kent State University, 141 p.

Sanders, L. L., 1991, Geochemistry of formation waters from the Lower Silurian Clinton Formation (Albion Sandstone), eastern Ohio: AAPG Bulletin, v. 75, p. 1593-1608.

Shindel, H.S., Stewart, L.J., and Kolva, J.R., 1983, Water Resources Data, Ohio Water Year 1982--volume 2, St. Lawrence River Basin Statewide Project Data: U.S. Geological Survey Water-Data Report OH-82-2, p. 277-286.

Shouakar-Stash, O. Evaluation of Stable Chlorine and Bromine Isotopes in Sedimentary Formation Fluids. (Ph.D. Thesis, University of Waterloo, 2008).

Siegel, D. I., Szustakowski, R. J., and Frape, S., 1990, Regional appraisal of brine chemistry in the Albion Group sandstones (Silurian) of New York, Pennsylvania, and Ohio: Association of Petroleum Geochemical Explorationists, Bulletin 6, p. 66-72.

Skeen, J. C., 2010, Basin Analysis and Aqueous Chemistry of Fluids in the Oriskany Sandstone, Appalachian Basin, USA: Master's Thesis, West Virginia University, Morgantown, WV, 109 p.

SPWLA, no date, RW Catalog database: Ohio Chapter of Society of Professional Well Log Analysts, unpublished.

Stueber, A.M., Pushkar, P., and Hetherington, E.A., 1987, A strontium isotopic study of formation waters from the Illinois basin, USA: Applied geochemistry, v. 2, no. 5, p. 477–494.

Stueber, A.M., Saller, A.H., and Ishida, H., 1998, Origin, migration, and mixing of brines in the Permian basin: geochemical evidence from the eastern central basin platform, Texas: AAPG bulletin, v. 82, no. 9, p. 1652–1672.

Stueber, A.M., and Walter, L.M., 1991, Origin and chemical evolution of formation waters from Silurian-Devonian strata in the Illinois basin, USA: Geochimica et Cosmochimica Acta, v. 55, no. 1, p. 309–325.

Stith, D.A., and Knapp, N.F., 1989, Characterization of trace metals in Ohio brines, final report: Ohio Division of Geological Survey Open-File Report 89-1, unpaginated.

Stith, D.A., Morth, A.H., and Hatch, J.R., 1979, Brine analyses, 1972-1974: Ohio Department of Natural Resources, Division of Geological Survey Open-File Report 79-1, 77 p.

Stout, Wilber, and Lamborn, R.E., 1924, The geology of Columbiana County: Ohio Division of Geological Survey Bulletin 28, 408 p.

Stout, W.E., Lamborn, R.E., and Schaaf, D., 1932, Brines of Ohio: Geological Survey of Ohio, Bulletin 37, 123 p.

Sunwall, M. T., and Pushkar, P., 1979, The isotopic composition of strontium in brines from petroleum fields of southeastern Ohio: Chemical Geology, v. 24, p. 189-197.

Thamke, J., 2014, written communication.

Thompson, W. E., 1973, Concentration of selected elements in brines of Perry County, Ohio: Ohio State University, Columbus Ohio, unpublished Master's thesis, 64 p.

U.S. Geological Survey National Oil and Gas Resource Assessment Team, 1995, 1995 National Assessment of United States Oil and Gas Resources: U.S. Geological Survey Circular 1118, 20 p.

Vugrinovich, 2013, written communication, Office of Oil, Gas, and Minerals, Michigan Department of Environmental Quality.

Walter, L.M., Budai, J.M., Martini, A., and Ku, T., 1997, Hydrogeochemistry of the Antrim Shale in the Michigan Basin. Final Report, Gas Research Institute, 5093-220-2704, 98 p.

Warner, N.R., Jackson, R.B., Darrah, T.H., Osborn, S.G., Down, A., Zhao, K., White, A., and Vengosh, A., 2012, Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania: PNAS, v. 109, no. 30, p. 11961–11966, Accessed www.pnas.org/cgi/doi/11910.11073/pnas.1121181109.

Wilk, G.B., 1987, Interpretation and geochemical analysis of "Clinton" brines from Mahoning County, Ohio: M.S. Thesis, University of Akron, 73 p.

Williams, J.H., Taylor, L.E., and Low, D.J., 1998, Hydrogeology and groundwater quality of the glaciated valleys of Bradford, Tioga, and Potter Counties, Pennsylvania: Pennsylvania Department of Natural Resources (DCNR); Pennsylvania Geological Survey, 4th Ser., Water Resource Report 68, 89 p.

Wyoming Oil and Gas Conservation Commission, Accessed 2013 at wogcc.state.wy.us.

Tables

Table 1. Input databases and datasets

[Short names of input databases, number of samples after removal of duplicates, and references on input databases. Please cite original references in future publications where there are known.]

ID of Database	Samples	Reference
ANTRIM	53	Walter and others (1997)
APPALACHIAN	1,600	Multiple – see references
ARKMOLDOVANYI	88	Moldovanyi and Walter (1992)
BARNABY	72	Barnaby and others (2004)
CAPO	58	Capo and others (2014)
CBM	3,220	Dahm (2013, writen communication)
CIMAREX	2,864	Cimarex Energy Company (2013, written communication)
FERRON	41	Rice (2003)
HOBBS	202	Hobbs and others (2011)
ILLINOIS	754	Meents and others (1952)
ILLINOISSTEUBER	90	Steuber and others (1987); Steuber and Walter (1991)
INDIANA	449	Keller (1983)
KHARAKA	316	Multiple – see references
KNOWLES	101	Powell and others (1963)
MAFLA	1541	Pashin (2013, written communication)
MICHIGAN	379	Vugrinovich (2013, written communication)
MISSISSIPPI	81	Carpenter and others (1974)
MONTANACBM	20	Meredith and others (2010)
NATCARB	57,116	Dept. of Energy, National Energy Technology Laboratory (2013)
NORTHDAKOTA	7,324	North Dakota Oil and Gas Division (2013)
OHBRINE	539	McDonald and others (2013, written communication)
PALODURO	16	Bassett and Bentley (1983)
PARADOX	91	Hanshaw and Hill (1969)
PASHIN	126	Alabama Geological Survey (2013, written communication)
PERMIAN	32	Steuber and others (1998)
POWDERRIVERCBM	47	Rice and others (2000)
ROCKIES	6,754	Dept. of Energy, National Energy Technology Laboratory (2005)
ROWANMARCELLUS	61	Rowan and others (2015)
SHOUAKARSTASH	126	Shouakar-Stash (2008)

SKEEN	352	Skeen (2010)
USGSBREIT	72,413	Breit (2002)
WILLISTON	48	Thamke (2015, written comm.), USGS OFRs 2010-1326 & 2012-1149
WYOGCC	10,020	Wyoming Oil and Gas Conservation Commission (2013)
Total	**166,994**	

Table 2. Variable names and descriptions

Variable Name	Description	Samples	Percent
IDUSGS	Unique ID in this database	166,994	100%
IDORIG	ID in original database or publication	166,994	100%
IDDB	ID (name) of input database	166,994	100%
SOURCE	Source of data	115,266	69%
REFERENCE	Publication	8,177	4.9%
LATITUDE	Latitude	155,153	93%
LONGITUDE	Longitude	155,146	93%
LATLONGAPX	Description if LATITUDE or LONGITUDE are approximate	22,955	14%
API	API well number, 14 digits	75,025	45%
USGSPROV	USGS Province	1,543	0.9%
USGSREGION	USGS Region	166,994	100%
BASIN	Basin	166,616	100%
BASINCODE	Basin Code	72,115	43%
STATE	State	166,994	100%
STATECODE	State Code	58,705	35%
COUNTY	County	85,315	51%
COUNTYCODE	County Code	130,826	78%
FIELD	Field	99,193	59%
FIELDCODE	Field Code	59,104	35%
WELLNAME	Well name	151,520	91%
WELLCODE	Well Code	18,602	11%
WELLTYPE	Well type	166,994	100%
WELLCLASS	Well class	45,153	27%
TOWNRANGE	Township, Range, Section, Quarter	129,344	77%
REGDIST	Regional District	20,099	12%
LOC	Location	3,979	2.4%
QUAD	Quad	137	0.1%
TIMESERIES	Order of time-series data	57	0.0%
DAY	Sample day of time-series data	2,806	1.7%
DATECOMP	Date of well completion	7,602	4.6%
DATESAMPLE	Date of sample collection	135,122	81%
DATEANALYS	Date of analysis	3,950	2.4%
METHOD	Sample Method	73,233	44%
OPERATOR	Well operator	35,107	21%

PERMIT	Well permit holder	8,339	5.0%
DFORM	Geologic formation name of greatest depth	24,397	15%
GROUP	Geologic group name	88	0.1%
FORMATION	Geologic formation name	113,575	68%
MEMBER	Geologic member name	2,586	1.5%
AGECODE	Geologic Age code	36,142	22%
ERA	Geologic Era name	28,853	17%
PERIOD	Geologic Period name	28,691	17%
EPOCH	Geologic Epoch name	23,379	14%
DEPTHUPPER	Upper perforation depth, ft. Depth added here if non-specific.	84,090	50%
DEPTHLOWER	Lower perforation depth, ft	102,389	61%
DEPTHWELL	Reported Total depth of well, ft	43,982	26%
ELEVATION	Elevation of well, ft	22,500	13%
LAB	Laboratory that analyzed the results	14,468	8.7%
REMARKS	Remarks or comments	54,423	33%
LITHOLOGY	Lithology	29,536	18%
POROSITY	Porosity, % reported	26	0.0%
TEMP	Temperature, deg F reported	4,495	2.7%
PRESSURE	Pressure, psi reported	907	0.5%
SG	Specific Gravity, reported or calculated (see text)	103,511	62%
SPGRAV	Specific Gravity, reported	64,746	39%
SPGRAVT	Temperature of Specific Gravity measurement, deg F	66,092	40%
RESIS	Resistivity, Ohm m	67,194	40%
RESIST	Temperature of Resistivity measurement, deg F	93,790	56%
PH	pH	123,129	74%
PHT	Temperature of pH measurement, deg F	37,349	22%
EHORP	Eh / Oxidation Reduction Potential, mV	31	0.0%
COND	Conductivity, µS/cm	1,024	0.6%
CONDT	Temperature of Conductivity measurement, deg F	35,237	21%
TURBIDITY	Turbidity	97	0.1%
HEM	Oil and Grease	130	0.1%
MBAS	Surfactants and Detergents	89	0.1%
UNITS	mg/L or ppm, applies to all chemistry unless specified	166,994	100%
TDSUSGS	Total Dissolved Solids, calculated (see text)	135,515	81%
TDS	Total Dissolved Solids, measured	101,019	60%
TDSCALC	Total Dissolved Solids, calculated, as reported in reference	1,988	1.2%

TSS	Total Suspended Solids	1,497	0.9%
CHARGEBAL	Charge balance of major ions, %, reported	7,052	4.2%
chargebalance	Charge balance of major ions, %, calculated	137,246	82%
Ag	Silver	307	0.2%
Al	Aluminum	2,353	1.4%
As	Arsenic	7,023	4.2%
Au	Gold	2	0.0%
B	Boron	16,147	10%
BO3	Borate	242	0.1%
Ba	Barium	15,976	10%
Be	Beryllium	369	0.2%
Bi	Bismuth	41	0.0%
Br	Bromide	8,914	5.3%
CO2	Carbon dioxide	1,766	1.1%
CO3	Carbonate	11,959	7.2%
HCO3	Bicarbonate	107,668	64%
Ca	Calcium	134,865	81%
Cd	Cadmium	819	0.5%
Ce	Cerium	1	0.0%
Cl	Chloride	141,122	85%
Co	Cobalt	640	0.4%
Cr	Chromium	4,450	2.7%
Cs	Cesium	427	0.3%
Cu	Copper	3,424	2.1%
F	Fluoride	18,258	11%
FeTot	Iron, total	36,897	22%
FeIII	Iron, 3+	476	0.3%
FeII	Iron, 2+	643	0.4%
FeS	Iron sulfide	1	0.0%
FeAl	Iron plus Aluminum, reported as elements	202	0.1%
FeAl2O3	Iron plus Aluminum, reported as oxides	444	0.3%
Ga	Gallium	88	0.1%
Ge	Germanium	8	0.0%
Hg	Mercury	781	0.5%
Hf	Hafnium	9	0.0%
I	Iodine	4,175	2.5%

Ir	Iridium	3	0.0%
K	Potassium	52,388	31%
KNa	Potassium plus Sodium	8,699	5%
La	Lanthanum	3	0.0%
Li	Lithium	9,842	5.9%
Mg	Magnesium	130,194	78%
Mn	Mangansese	10,023	6.0%
Mo	Molybdenum	1,778	1.1%
N	Nitrogen, total	242	0.1%
NO2	Nitrite	105	0.1%
NO3	Nitrate	4,641	2.8%
NO3NO2	Nitrate plus Nitrite	103	0.1%
NH3	Ammonia	324	0.2%
NH4	Ammonium	1,414	0.8%
TKN	Kjeldahl Nitrogen	134	0.1%
Na	Sodium	113,390	68%
Nb	Niobium	10	0.0%
Ni	Nickel	1,383	0.8%
DO	Dissolved Oxygen	23	0.0%
OH	Hydroxide	444	0.3%
Os	Osmium	2	0.0%
P	Phosphorus	24	0.0%
PO4	Phosphate	418	0.3%
Pb	Lead	1,105	0.7%
Re	Rhenium	1	0.0%
Rh	Rhodium	1	0.0%
Rb	Rubidium	1,025	0.6%
Ru	Ruthenium	6	0.0%
S	Sulfide	170	0.1%
SO3	Sulfite	80	0.0%
SO4	Sulfate	121,323	73%
HS	Bisulfide	20	0.0%
H2S	Hydrogen Sulfide	3,401	2.0%
Sb	Antimony	420	0.3%
Sc	Scandium	43	0.0%
Se	Selenium	1,927	1.2%

Si	Silica	3,575	2.1%
Sn	Tin	379	0.2%
Sr	Strontium	11,014	6.6%
Ta	Tantalum	1	0.0%
Te	Tellurium	2	0.0%
Th	Thorium	90	0.1%
Ti	Titanium	483	0.3%
Tl	Thallium	96	0.1%
U	Uranium	23	0.0%
V	Vanadium	1,168	0.7%
W	Tungsten	67	0.0%
Zn	Zinc	4,772	2.9%
Zr	Zirconium	281	0.2%
ALKCACO3	Alkalinity as CaCO3	5,047	3.0%
ALKHCO3	Alkalinity as HCO3	1,156	0.7%
ACIDITY	Acidity as CaCO3	102	0.1%
DIC	Dissolved Inorganic Carbon	99	0.1%
DOC	Dissolved Organic Carbon	175	0.1%
TOC	Total Organic Carbon	433	0.3%
CN	Cyanide	112	0.1%
BOD	Biochemical Oxygen Demand	114	0.1%
COD	Chemical Oxygen Demand	115	0.1%
ACETATE	Acetate	129	0.1%
ACETICACID	Acetic Acid	960	0.6%
BENZENE	Benzene	934	0.6%
PHENOLS	Phenols	109	0.1%
TOLUENE	Toluene	896	0.5%
XYLENE	Xylene	39	0.0%
ETHYLBENZ	Ethybenzene	7	0.0%
ALPHA	Alpha particle (^{4}He), pCi/L	64	0.0%
BETA	Beta particle, pCi/L	73	0.0%
dD	δH, per mil	1,190	0.7%
H3	Tritium, ^{3}H, tritium units	36	0.0%
d13C	δδ^{13}C, per mil	78	0.0%
C14	^{14}C, pCi/L	9	0.0%
d18O	δδ^{18}O, per mil	1,192	0.7%

d34S	$\delta\delta^{34}S$, per mil	4	0.0%
d37Cl	$\delta\delta^{37}Cl$, per mil	259	0.2%
K40	^{40}K, pCi/L	50	0.0%
d81Br	$\delta^{81}Br$	94	0.1%
Sr87Sr86	$^{87}Sr/^{86}Sr$	716	0.4%
I129	^{129}I, pCi/L	12	0.0%
Rn222	^{222}Rn, pCi/L	134	0.1%
Ra226	^{226}Ra, pCi/L	696	0.4%
Ra228	^{228}Ra, pCi/L	158	0.1%
cull_PH	"X" if pH < 4.5 or pH > 10.5		
cull_MgCa	"X" if Mg > Ca		
cull_KCl	"X" if K > Cl		
cull_K5Na	"X" if K > 5xNa		
cull_chargeb	"X" if charge balance > 15%		